

# Decoding square-free Goppa codes over $\mathbb{F}_p$

Paulo S. L. M. Barreto, Rafael Misoczki, and Richard Lindner

**Abstract**—We propose a new, efficient non-deterministic decoding algorithm for square-free Goppa codes over  $\mathbb{F}_p$  for any prime  $p$ . If the code in question has degree  $t$  and the average distance to the closest codeword is at least  $(4/p)t + 1$ , the proposed decoder can uniquely correct up to  $(2/p)t$  errors with high probability. The correction capability is higher if the distribution of error magnitudes is not uniform, approaching or reaching  $t$  errors when any particular error value occurs much more often than others or exclusively. This makes the method interesting for (semantically secure) cryptosystems based on the decoding problem for permuted and punctured Goppa codes.

**Index Terms**—Algorithms, Cryptography, Decoding, Error correction

## I. INTRODUCTION

**P**UBLIC-KEY cryptosystems based on coding theory, known for nearly as long as the very concept of asymmetric cryptography itself, have recently been attracting renewed interest because of their apparent resistance even against attacks mounted with the help of quantum computers, constituting a family of so-called post-quantum cryptosystems [1]. However, not all error-correcting codes are suitable for cryptographic applications. The most commonly used family of codes for such purposes is that of Goppa codes, which remain essentially unharmed by cryptanalysis efforts despite considerable efforts and progress in the area.

Introduced in 1970, Goppa codes [2] are a subfamily of alternant codes, i.e. subfield subcodes of Generalized Reed-Solomon codes. Let  $q = p^m$  for some prime  $p$  and some  $m > 0$ . A Goppa code  $\Gamma(L, g)$  over  $\mathbb{F}_p$  is determined by a sequence  $L \in \mathbb{F}_q^n$  of distinct values, and a polynomial  $g \in \mathbb{F}_q[x]$  of degree  $t := \deg(g)$  whose roots are disjoint from  $L$ . Goppa codes have by design a minimal distance at least  $t + 1$  by virtue of being alternant. Certain codes are known to have better minimum distances than this lower bound. Thus, binary Goppa codes where  $g$  is square-free are known to have a larger minimum distance of at least  $2t + 1$  instead. A family of codes where  $g$  is not square-free have minimum distance at least  $t + \gamma - 1$  for some  $2 < \gamma < t - 1$ , which is known as the Hartmann-Tzeng bound for Goppa codes [3], [4].

The class of Sugiyama-Kasahara-Hirasawa-Namekawa codes [5] where  $g = h^{r-1}$  for some square-free monic poly-

nomial  $h \in \mathbb{F}_q[x]$  and some power  $r$  of  $p$  dividing  $q$ , which constitute a proper superclass of the so-called “wild” codes where  $h$  is restricted to being irreducible [6], have minimum distance at least  $r \deg(h) + 1$  rather than  $(r - 1) \deg(h) + 1$ . Although it is known that the minimum distance of a Goppa code of degree  $t$  is at least  $t + 1$  and there are known cases where it is higher (up to  $2t + 1$ , as it happens for binary square-free Goppa codes), systematically determining the true minimum distance of any given subfamily of Goppa codes remains largely an open problem, yet it is an important metric as it determines not only how many errors can always be uniquely corrected, but indirectly the security level and the key sizes of the cryptosystems based on each given code.

Apart from brute force, known decoding methods for alternant codes can in general correct only about half as many errors as a binary square-free Goppa code is in principle able to correct [7], [8] (see also [9]). Even the Guruswami-Sudan algorithm [10], which exceeds the  $t/2$  limit, can only correct about  $n - \sqrt{n(n-t)} \approx t/2 + (t/2)^2/(2n-t)$  errors. In contrast, Patterson’s algorithm can correct all  $t$  design errors of binary Goppa codes, as can an alternant decoder using the equivalence  $\Gamma(L, g) = \Gamma(L, g^2)$  albeit at a larger computational cost. Bernstein’s list decoding method [11] goes somewhat further, attaining a correction capability of  $n - \sqrt{n(n-2t-2)} \approx t + 1 + (t+1)^2/(2n-t-1)$  errors for binary irreducible Goppa codes, although decoding is ambiguous if the actual distance is not proportionally higher. Similar techniques can in principle correct about  $n - \sqrt{n(n-rt)} \approx rt/2 + (rt/2)^2/(2n-rt)$  errors for wild codes [6]. Bernstein’s method does not reach the  $q$ -ary Johnson radius, but a more recent algorithm by Augot *et al.* does so in the binary case [12].

## A. Our Results

Our contribution in this paper is a non-deterministic decoding algorithm for square-free Goppa codes over  $\mathbb{F}_p$  for any prime  $p$ . The method generalizes Patterson’s approach and can potentially correct up to  $(2/p)t$  errors, on the condition that a suitable short vector can be found in a certain polynomial lattice. In particular, our method corrects  $(2/3)t$  errors in characteristic 3, exceeding the  $t/2$  barrier when the average distance to the closest codeword is at least  $(4/3)t + 1$ . In experiments conducted to assess the practical behaviour of our proposal, the result of the decoding is observed to be unique with overwhelming probability for irreducible ternary Goppa codes chosen uniformly at random, hinting that, for the vast majority of such codes, the average distance to the closest codeword is sufficiently higher than the ensured minimum distance. Besides, our proposal can probabilistically correct a still larger number of errors that approaches and

P. Barreto is supported by the Brazilian National Council for Scientific and Technological Development (CNPq) under research productivity grant 303163/2009-7.

P. Barreto is with the Department of Computer and Digital Systems Engineering, Escola Politécnica, Universidade de São Paulo, Brazil. (e-mail: pbarreto@larc.usp.br)

R. Misoczki is with SECRET team, INRIA Paris-Rocquencourt, France. (e-mail: rafael.misoczki@inria.fr)

R. Lindner is with the Department of Computer Science, Technische Universität Darmstadt, Darmstadt, Germany. (e-mail: rlindner@cdc.informatik.tu-darmstadt.de)

reaches  $t$  depending on the distribution of error magnitudes. For instance, the method corrects up to  $t$  errors with high probability if all error magnitudes are known to be equal.

This feature outperforms even Sugiyama-Kasahara-Hirasawa-Namekawa and wild codes and the associated decoding methods, and is particularly interesting for cryptographic applications like McEliece encryption [13] under the Fujisaki-Okamoto or similar semantic security transform [14], where error magnitudes can be chosen by convention to be all equal. In that case, even if an attacker could somehow derive a generic alternant decoder from the public code that is typical in such systems (a strategy exploited e.g. in [15]), he will not be able to correct more than about  $t/2$  errors out of roughly  $t$  that can be corrected with the private trapdoor enabled by our proposal, facing an infeasible workload of about  $(p-1)\binom{n}{t/2}/\binom{t}{t/2}$  guesses to mount a complete attack. This makes Goppa codes in odd characteristic, which have already been shown to sport some potential security advantages over binary ones [16], even more attractive in practice.

For the benefit of implementors, we describe a dedicated version of the Mulders-Storjohann algorithm to convert the particular lattice basis encountered during the decoding process to weak Popov form. The computational complexity of this step is then shown to be  $O(p^3 t^2)$ .

## B. Organization of the Paper

The remainder of this document is organized as follows. We provide basic notions in Section II. We recapitulate Patterson's decoding algorithm for binary irreducible Goppa codes in Section III, and extend it to square-free codes in characteristic  $p$  in Section IV, showing that it can correct  $(2/p)t$  errors in general and up to  $t$  errors depending on the distribution of error magnitudes. We conclude in Section V.

## II. PRELIMINARIES

Matrix indices will start from 0 throughout this paper, unless otherwise stated. Let  $p$  be a prime and let  $q = p^m$  for some  $m > 0$ . The finite field of  $q$  elements is written  $\mathbb{F}_q$ . For sequences of elements  $(g_1, \dots, g_t) \in \mathbb{F}_q^t$ ,  $(L_0, \dots, L_{n-1}) \in \mathbb{F}_q^n$  and  $(d_0, \dots, d_{n-1}) \in \mathbb{F}_q^n$  for some  $t, n \in \mathbb{N}$ , we denote by  $\text{toep}(g_1, \dots, g_t)$  the  $t \times t$  Toeplitz matrix with elements  $T_{ij} := g_{t-i+j}$  for  $j \leq i$  and  $T_{ij} := 0$  otherwise; by  $\text{vdm}_t(L_0, \dots, L_{n-1})$  the  $t \times n$  Vandermonde matrix with elements  $V_{ij} := L_j^i$ ,  $0 \leq i < t$ ,  $0 \leq j < n$ ; and by  $\text{diag}(d_0, \dots, d_{n-1})$  the diagonal matrix with diagonal elements  $D_{jj} := d_j$ ,  $0 \leq j < n$ .

### A. Error Correcting Codes

Let  $L = (L_0, \dots, L_{n-1}) \in \mathbb{F}_q^n$  be a sequence (called the *support*) of  $n \leq q$  distinct elements, and let  $g \in \mathbb{F}_q[x]$  be an irreducible monic polynomial of degree  $t$  such that  $g(L_i) \neq 0$  for all  $i$ . For any word  $e \in \mathbb{F}_p^n$  we define the corresponding *Goppa syndrome* polynomial  $s_e \in \mathbb{F}_q[x]$  to be:

$$s_e(x) = \sum_{i=0}^{n-1} \frac{e_i}{x - L_i} \mod g(x).$$

Thus the syndrome is a linear function of  $e$ . The  $[n, \geq n - mt, \geq t + 1]$  Goppa code over  $\mathbb{F}_p$  with support  $L$  and generator polynomial  $g$  is the kernel of the syndrome function applied to elements from  $\mathbb{F}_p^n$ , i.e. the set  $\Gamma(L, g) := \{e \in \mathbb{F}_p^n \mid s_e \equiv 0 \mod g\}$ .

Writing  $s_e(x) := \sum_i s_i x^i$  for some  $s \in \mathbb{F}_q^t$ , one can show that  $s^T = H e^T$  where the *parity-check matrix*  $H$  has the form

$$\begin{aligned} H &= \begin{bmatrix} \text{toep}(g_1, \dots, g_t) \\ \vdots \\ \text{vdm}_t(L_0, \dots, L_{n-1}) \\ \vdots \\ \text{diag}(g(L_0)^{-1}, \dots, g(L_{n-1})^{-1}) \end{bmatrix} \end{aligned} \quad (1)$$

Thus  $H = TVD \in \mathbb{F}_q^{t \times n}$ , where  $T \in \mathbb{F}_q^{t \times t}$  is a Toeplitz matrix,  $V \in \mathbb{F}_q^{t \times n}$  is a Vandermonde matrix, and  $D \in \mathbb{F}_q^{n \times n}$  is a diagonal matrix.

Since a Goppa code is a  $\mathbb{F}_p$ -subfield subcode, it is possible to express the syndrome function in terms of a parity-check matrix  $\bar{H} \in \mathbb{F}_p^{mt \times n}$  using the so-called *trace construction* (see e.g. [9, Ch. 7, § 7]). This is useful to obtain a syndrome  $\bar{s} \in \mathbb{F}_p^{mt}$  equivalent to  $s \in \mathbb{F}_q^t$  above while keeping the arithmetic operations in  $\mathbb{F}_p$  rather than  $\mathbb{F}_q$ , even though it is not immediately useful for decoding, at which point a syndrome over  $\mathbb{F}_q$  has to be assembled by inverting the trace construction.

The *syndrome decoding problem* consists of computing the error pattern  $e$  given its syndrome  $s_e$ . Knowledge of the code structure in the form of the support  $L$  and the polynomial  $g$  makes this problem solvable in polynomial time, with some constraints relating the weight of  $e$  to the degree of  $g$ .

### B. Polynomial Lattices

Let  $A \in \mathbb{F}_q[x]^{n \times m}$  be a polynomial matrix, and let  $r$  denote its rank (i.e. assume that  $A$  has  $r$  linearly independent rows). The (polynomial) lattice  $\Lambda(A)$  over  $\mathbb{F}_q[x]$  spanned by the rows of  $A$  is

$$\Lambda(A) = \{(u_0, \dots, u_{n-1})A \in \mathbb{F}_q[x]^m \mid (u_0, \dots, u_{n-1}) \in \mathbb{F}_q[x]^n\}.$$

The notion of length which we will use for  $f \in \mathbb{F}_q[x]$  is  $|f| = \deg(f)$ . For polynomial vectors  $v \in \mathbb{F}_q[x]^n$  we adopt the notion of maximal degree length:  $|v| = \max_i |v_i|$ . This notion is coarse enough that, contrary to integer lattices where finding even an approximation to the shortest vector by a constant factor is a hard problem [17], reducing a basis for a polynomial lattice can be achieved in polynomial time. The following result by Mulders and Storjohann holds [18]:

**Theorem 1.** *There exists an algorithm which finds the shortest nonzero vector in the  $\mathbb{F}_q[x]$ -module generated by the rows of  $A$  with  $O(mnrd^2)$  operations in  $\mathbb{F}_q$ , where  $d = \max\{\deg(A_{ij}) \mid 1 \leq i \leq n, \leq j \leq m\}$ .*

The algorithm whose existence is established by Theorem 1 is based on converting a given lattice basis to the so-called weak Popov form, formally defined in Appendix A, which also contains a description of Algorithm 2 and its cost behavior in the context of decoding.

The weak Popov form is not the only way to find short vectors in a polynomial lattice, and in fact this is not critical to our proposal in this paper; for instance, the method by Lee and O'Sullivan [19], which is related to Gröbner bases, would appear to be an alternative. Our choice of the Mulders-Storjohann

method, which computes the weak Popov form, derives from its conceptual simplicity and ease of implementation, since the result turns out to be a natural generalization of Patterson's decoding algorithm described in the next section.

### III. PATTERSON'S DECODING METHOD

We briefly recapitulate Patterson's decoding algorithm [20], which will provide the basis for the general algorithm we propose. The goal, of course, is to compute the error pattern  $e$  given its syndrome  $s_e$  and the structure of  $\Gamma(L, g)$ .

Let  $q = 2^m$ , and assume we are given a binary Goppa code  $\Gamma(L, g)$  where the monic polynomial  $g$  is irreducible. We define the Patterson locator polynomial  $\sigma \in \mathbb{F}_q[x]$  as:

$$\sigma(x) := \prod_{e_i=1} (x - L_i). \quad (2)$$

The name *locator polynomial* comes from the fact that the roots of  $\sigma$  clearly indicate where errors occurred, since  $\sigma(L_j) = 0 \Leftrightarrow e_j = 1$ . Taking the derivative of the formal power series underlying  $\sigma$ , we obtain

$$\begin{aligned} \sigma'(x) &= \sum_{e_i=1} \prod_{\substack{e_j=1 \\ j \neq i}} (x - L_j) \\ &= \sum_{e_i=1} \frac{1}{x - L_i} \prod_{e_j=1} (x - L_j) \\ &= \sigma(x) \sum_i \frac{e_i}{x - L_i}, \end{aligned}$$

and hence, in  $\mathbb{F}_q[x]/g(x)$ ,

$$\sigma'(x) = \sigma(x)s_e(x) \pmod{g(x)}. \quad (3)$$

This is called the *key equation*, and now we discuss how to solve it.

Being a polynomial in characteristic 2,  $\sigma(x)$  modulo  $g(x)$  can be written as

$$\sigma(x) = a_0(x)^2 + xa_1(x)^2$$

for some  $a_0(x)$ ,  $a_1(x)$  with  $\deg(a_0) \leq \lfloor t/2 \rfloor$  and  $\deg(a_1) \leq \lfloor (t-1)/2 \rfloor$ , and hence

$$\sigma'(x) = 2a_0(x)a_0'(x) + a_1(x)^2 + 2a_1(x)a_1'(x)x = a_1(x)^2,$$

since the characteristic is 2. Therefore

$$\begin{aligned} a_1(x)^2 &= \sigma'(x) = \sigma(x)s_e(x) \\ &= (a_0(x)^2 + xa_1(x)^2)s_e(x) \pmod{g(x)}, \end{aligned}$$

whence

$$a_0(x) = a_1(x)v(x) \pmod{g(x)} \quad (4)$$

where  $v(x)$  is a polynomial satisfying  $v(x)^2 = x + 1/s_e(x) \pmod{g(x)}$ . Such a polynomial surely exists in characteristic 2 if  $g(x)$  is square-free: if  $g(x) = \prod_i g_i(x)$  where each  $g_i(x)$  is irreducible, then  $v(x) \pmod{g_i(x)}$  can be computed as a square root of  $x + 1/s_e(x) \pmod{g_i(x)}$  viewed as an element of the finite field  $\mathbb{F}_q[x]/g_i(x)$ , and  $v(x) \pmod{g(x)}$  can then be obtained by combining the results via the Chinese Remainder Theorem. We can thus assume that  $\deg(v) < t = \deg(g)$ .

The last equation is actually a Bézout relation  $a_0(x) = a_1(x)v(x) + \lambda(x)g(x)$ , which can be solved for  $a_0(x)$  and  $a_1(x)$  with the restriction  $\deg(a_j) \leq \lfloor (t-j)/2 \rfloor$  (and also  $\lambda(x)$  but it is not used) using the extended Euclidean algorithm. Solutions  $(a_0, a_1)$  can also be seen as short vectors in the lattice spanned by the rows of the following matrix:

$$A = \begin{bmatrix} g & 0 \\ v & 1 \end{bmatrix}$$

in the sense that the degrees of these polynomials are much smaller than uniformly random vectors, since  $(\lambda, a_1)A = (\lambda g + a_1 v, a_1) = (a_0, a_1)$  for some  $\lambda \in \mathbb{F}_q[x]$ , by virtue of Equation 4. Therefore, Algorithm 2 can be used to find candidate solutions  $(a_0, a_1)$ .

At first glance there is no guarantee that a short vector in the lattice generated by  $A$  yields the desired solution; in other words, being short is a necessary condition, but in principle not a sufficient one. However, the fact that in the binary case the minimum code distance is known [2] to be at least  $2t + 1$  actually restricts  $\sigma$  to a single candidate, so that Algorithm 2 is bound to find it. Thus, decoding is always successful up to  $t$  introduced errors.

### IV. DECODING CODES OVER $\mathbb{F}_p$

We now show how to generalize Patterson's decoding algorithm so as to correct errors for codes defined over general prime fields. Thus, let  $q = p^m$  for some prime  $p$  and some  $m > 0$ , and assume we are given an irreducible Goppa code  $\Gamma(L, g)$  over  $\mathbb{F}_p$ .

Let  $\phi \in \mathbb{F}_p \setminus \{0\}$  be a constant scalar. We define the generalized error locator polynomial to be

$$\sigma_\phi(x) := \prod_i (x - L_i)^{e_i/\phi} \quad (5)$$

where the value  $e_i/\phi$  is lifted from  $\mathbb{F}_p$  to  $\mathbb{Z}$  (i.e. the value  $e_i/\phi \in \mathbb{F}_p$  that occurs as an exponent is taken to mean its corresponding least non-negative integer representative, which lies in range  $0 \dots p-1$ ). One can easily see that this definition actually coincides with Patterson error locator polynomials as defined by Equation 2 for  $p = 2$ . Lifting Equation 5 to the field of rational functions in characteristic 0 and taking the derivative, we have

$$\begin{aligned} \sigma'_\phi(x) &= \sum_j (e_j/\phi)(x - L_j)^{e_j/\phi-1} \prod_{i \neq j} (x - L_i)^{e_i/\phi} \\ &= (1/\phi) \sum_j \frac{e_j}{x - L_j} \prod_i (x - L_i)^{e_i/\phi} \\ &= (1/\phi) \sigma_\phi(x) \sum_j \frac{e_j}{x - L_j}, \end{aligned}$$

which over  $\mathbb{F}_q[x]$  reduces to

$$\phi \sigma'_\phi(x) = \sigma_\phi(x)s_e(x) \pmod{g(x)}. \quad (6)$$

This is the  $\phi$ -th key equation of the proposed method, which generalizes Equation 3 to Goppa codes over  $\mathbb{F}_p$ . The actual  $\phi$  must be chosen so as to minimize the degree of  $\sigma_\phi$  (and hence maximize the number of correctable errors). One cannot expect to know *a priori* the value of  $\phi$ , but since there are only  $p-1$

possibilities, the error correction strategy will be to try each of them in turn.

Notice that the maximum number of correctable errors can be, and usually is, less than  $t$ , since the degree of  $\sigma_\phi$  exceeds the number of roots in the presence of multiple roots. The following theorem provides an upper bound for how many errors can be corrected by solving Equation 6 when the distribution of error magnitudes is not taken into account.

**Theorem 2.** *The maximum number of errors that can be corrected by solving Equation 6 independently of the distribution of error magnitudes is  $w = (2/p)t$ .*

*Proof:* Let  $w_v$  denote the number of times the magnitude  $v$  occurs in an error pattern of weight  $w$ , so that  $\sum_v w_v = w$ . Since we are working with a Goppa code, the constraint for correctability is  $\deg(\sigma_\phi) = \sum_v (v/\phi)w_v \leq t$ . In the extreme situation when the weight of the error pattern reaches  $w$ , the most often error magnitude occurs  $w_{\max} \geq w/(p-1)$  times, attaining the lower bound when all error magnitudes occur with equal frequency. In that case,  $\deg(\sigma_\phi) \leq \sum_v (v/\phi)w_{\max} = (1+2+\dots+(p-1))w/(p-1) = wp/2 \leq t$ , and hence no more than errors than  $w = (2/p)t$  can be corrected independently of the distribution of magnitudes, as claimed. ■

Since the proposed method coincides with Patterson's for  $p = 2$ , it is not surprising that  $t$  errors can be corrected in that case. However, in characteristic 3 the number of potentially correctable errors is  $(2/3)t$ , non-deterministically exceeding the limit of  $t/2$  errors attainable by previously known decoding methods for codes of degree  $t$ , except in the case of so called "wild codes" [6] whereby the Goppa polynomial is a  $(p-1)$ -th power of an irreducible polynomial (our method, by contrast, applies when that polynomial is square-free, as we will see in Section IV-A).

Despite the low general limit of  $(2/p)t$  correctable errors for  $p > 3$ , it is still possible to exceed that limit in any odd characteristic if the distribution of error magnitudes is unbalanced. Indeed, all that is required to get a chance of uniquely decoding a word containing  $w \leq t$  errors is that  $\deg(\sigma_\phi) \leq t$  for some choice of  $\phi$  and that the actual distance from the right codeword to any other codeword be at least  $2w + 1$ .

The actual number of correctable errors depends heavily on the distribution of error magnitudes and has to be computed in a case-by-case basis, always laying in the range  $(2/p)t$  to  $t$ . In particular, if all error magnitudes are equal, in principle one could correct  $t$  errors, even though this is a statistical rather than deterministic behavior. This is especially useful for cryptographic applications involving an all-or-nothing transform [21], as it happens e.g. for a semantically secure encryption scheme involving the McEliece one-way trapdoor function [13] and the Fujisaki-Okamoto conversion [14]. In such scenarios, the magnitudes of the introduced errors can be chosen to be all or nearly all equal by convention, making the proposed decoder attractive for its higher decodability bound, under the explicit assumption that decoding them remains hard.

#### A. Solving the Key Equation

We now focus on actually solving Equation 6. Being a polynomial in characteristic  $p$ ,  $\sigma_\phi(x)$  can be written as

$$\sigma_\phi(x) = \sum_{k=0}^{p-1} x^k a_k(x)^p \quad (7)$$

for some  $a_k(x)$  with  $\deg(a_k) \leq \lfloor (t-k)/p \rfloor$ ,  $0 \leq k \leq p-1$ , and hence

$$\begin{aligned} \sigma'_\phi(x) &= \sum_{k=0}^{p-1} (kx^{k-1} a_k(x)^p + px^k a_k(x)^{p-1} a'_k(x)) \\ &= \sum_{k=1}^{p-1} kx^{k-1} a_k(x)^p \end{aligned}$$

since the characteristic is  $p$ . Therefore

$$\begin{aligned} \phi \sum_{k=1}^{p-1} kx^{k-1} a_k(x)^p &= \phi \sigma'_\phi(x) = \sigma_\phi(x) s_e(x) \\ &= \left( \sum_{k=0}^{p-1} x^k a_k(x)^p \right) s_e(x) \bmod g(x), \end{aligned}$$

whence

$$a_0 + \sum_{k=1}^{p-1} a_k(x) v_k(x) = 0 \bmod g(x) \quad (8)$$

where the  $v_k(x)$  are polynomials satisfying  $v_k(x)^p = x^k - \phi k x^{k-1} / s_e(x) \bmod g(x)$ . Such polynomials surely exist in characteristic  $p$  if  $g(x)$  is square-free: if  $g(x) = \prod_i g_i(x)$  where each  $g_i(x)$  is irreducible, then  $v_k(x) \bmod g_i(x)$  can be computed as a  $p$ -th root of  $x^k - \phi k x^{k-1} / s_e(x) \bmod g_i(x)$  viewed as an element of the finite field  $\mathbb{F}_q[x]/g_i(x)$ , and  $v_k(x) \bmod g(x)$  can then be obtained by combining the results via the Chinese Remainder Theorem. We can thus assume that  $\deg(v_k) < t = \deg(g)$ .

The Diophantine equation 8 has to be solved for  $a_k(x)$  with the stated restriction on their degrees. Solutions  $(a_0, a_1, \dots, a_{p-1})$  can be seen as short vectors in the lattice spanned by the rows of the matrix

$$A_\phi = \begin{bmatrix} g & 0 & 0 & \dots & 0 \\ -v_1 & 1 & 0 & \dots & 0 \\ -v_2 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -v_{p-1} & 0 & 0 & \dots & 1 \end{bmatrix}, \quad (9)$$

since, by virtue of Equation 8, one has  $(\lambda, a_1, \dots, a_{p-1}) A_\phi = (\lambda g - \sum_{k=1}^{p-1} a_k(x) v_k(x), a_1, \dots, a_{p-1}) = (a_0, a_1, \dots, a_{p-1})$  for some  $\lambda \in \mathbb{F}_q[x]$ . Therefore, Algorithm 2 can be used to find candidate solutions  $(a_0, \dots, a_{p-1})$ .

The method is applicable whenever one can actually invert  $s \bmod g$  and then compute the  $p$ -roots needed to define the  $v_k$  polynomials. This is always the case when  $g$  is irreducible, but not exclusively so. Indeed, to compute the  $v_k$  it suffices that  $g$  is square-free and that  $s$  is invertible modulo each of the irreducible factors of  $g$ , since in this case the  $v_k$  can be easily computed modulo those irreducible factors and finally recovered via the Chinese Remainder Theorem.

Theorem 1 ensures a cost not exceeding  $O(p^3 t^2)$   $\mathbb{F}_q$  operations for computing short vectors in  $\Lambda(A_\phi)$ .

### B. Estimating the Success Probability

Regrettably the ability to find short vectors in lattice  $\Lambda(A_\phi)$  does not mean that any such vector yields a solution to Equation 6. We will now see that, fortunately, the proposed method has a surprisingly favourable probability of finding the right  $(a_0, \dots, a_{p-1})$  that solves Equation 6.

As a cautionary note, we stress that a full theoretical analysis of the failure probability remains, to the best of our knowledge, an open problem. Because the distance notion here is not Euclidean, but rather that of Hamming, Minkowski's theorem on the existence of a lattice point in any large enough convex set does not appear to apply in our case. Also, the analysis would also appear to require a detailed theory on the distribution of the average distance between a vector and the closest codeword whose magnitudes satisfy some constraint (like being all equal, or following a highly skewed distribution), which to the best of our knowledge is too an open problem. As a consequence, the failure probability estimates we provide are conjectured on an empirical basis. Namely, they result from experiments conducted on a large number of random Goppa codes at which our decoding method is targeted, and for each of those codes, on a large number of decoding attempts on random error patterns following the particular magnitude constraint (equal for all error positions) for which the decoder works best.

In a successful decoding, the reduced basis for lattice  $\Lambda(A_\phi)$  leads to candidates for  $\sigma_\phi$  with degree  $t$  onward, of which of course only the candidate with the smallest degree is the correct one. Spurious candidates of degree close to  $t$  result from random-looking short (albeit not shortest) vectors in the reduced basis and are usually harmless. But the fact that those short vectors are “random-looking” means they are also a threat: if by chance they are such that the coefficient of the highest-degree term in the associated spurious  $\sigma_\phi$  vanishes,  $\deg(\sigma_\phi)$  becomes  $t$  or less. Since this is connected with the vanishing of a coefficient from  $\mathbb{F}_q$ , this event happens with probability  $1/q$  assuming that short spurious vectors in  $\Lambda(A_\phi)$  are approximately uniformly distributed.

In general, when trying to correct  $w < t$  errors of equal magnitude for a uniformly random irreducible Goppa code, the top  $t+1-w$  coefficients in the spurious  $\sigma_\phi$  must vanish to interfere with the decoding process, whence the probability of successful decoding is roughly  $1 - 1/q^{t+1-w}$ . This matches the empirically observed behaviour of the proposed method in odd characteristic. Not surprisingly, the method always succeeds for binary codes, since it reduces to Patterson's algorithm and the minimum code distance is known to be at least  $2t+1$ .

Table I illustrates the results of experiments in Magma [22] supporting the conjecture that the probability of successful decoding is roughly  $\Pr_{suc} := 1 - 1/q^{t+1-w}$ . For each quadruple  $(p, m, t, w)$ , a set of 10000 Goppa codes of maximum length  $n = p^m$  and degree  $t$  plus an error pattern of length  $n$ , weight  $w$  and all magnitudes equal (to a single random value in  $\mathbb{F}_p \setminus \{0\}$ ) were randomly generated, and the proposed method was then applied to decode the syndrome of that error pattern. The predicted number of successful decodings,  $N_{pre}$ , is then compared with the actually observed number  $N_{obs}$  of

successful decodings. For each combination of  $p$  and  $m$ , the first listed  $t$  is the largest integer satisfying  $p^m - mt > 0$ . The number  $w$  of introduced errors is then decreased starting from  $t$  until the probability of success exceeds 0.9999. Since  $\Pr_{suc}$  is close to 1 for large  $q$ , reasonably small values are chosen for all parameters so that the probability of decoding failure is large enough to be easily discernible. This is also the reason why more detail is provided for smaller parameters. We omit the results for characteristic 2, since in all tests we conducted no decoding failure was observed, as expected. We stress that these examples are not meant by any means for practical cryptographic use.

TABLE I  
EXPERIMENTAL ASSESSMENT OF THE PROBABILITY OF DECODING SUCCESS

$p$	$m$	$t$	$w$	$\Pr_{suc}$	$N_{pre}$	$N_{obs}$
3	3	8	8	0.962963	9630	9670
3	3	8	7	0.998628	9986	9992
3	3	8	6	0.999949	9999	9999
3	3	7	7	0.962963	9630	9639
3	3	7	6	0.998628	9986	9989
3	3	7	5	0.999949	9999	10000
3	3	6	6	0.962963	9630	9645
3	3	6	5	0.998628	9986	9991
3	3	6	4	0.999949	9999	10000
3	4	20	20	0.987654	9877	9883
3	4	20	19	0.999848	9998	9997
3	4	20	18	0.999998	10000	10000
5	2	12	12	0.960000	9600	9612
5	2	12	11	0.998400	9984	9985
5	2	12	10	0.999936	9999	10000
5	3	41	41	0.992000	9920	9924
5	3	41	40	0.999936	9999	10000
7	2	24	24	0.997085	9971	9989
7	2	24	23	0.999992	9999	10000
11	2	60	60	0.991736	9917	9922
11	2	60	59	0.999932	9999	9999

Decoding  $w \leq (2/p)t$  errors of uniformly random magnitude for a uniformly random irreducible code is of course always successful for  $p > 3$ , since in that case  $w$  does not exceed half the minimum code distance, which is at least  $t+1 > (4/p)t+1 \geq 2w+1$ .

### C. Computing the Error Magnitudes

In contrast to generic alternant decoding methods, there is no need to compute an error evaluator polynomial to obtain the error magnitudes in the current proposal. After obtaining  $\sigma_\phi(x)$  and finding its roots  $L_j$ , all that is needed to compute the corresponding error values  $e_j$  is to determine the multiplicity  $\mu_j$  of each root, since one can see from Equation 5 that  $e_j = \phi\mu_j$ .

Computing  $\mu_j$  is accomplished by determining how many times  $(x - L_j) \mid \sigma_\phi(x)$  whenever  $\sigma_\phi(L_j) = 0$ , or alternatively by finding the highest derivative of  $\sigma_\phi$  such that  $\sigma_\phi^{(h)}(L_j) = 0$  (and setting  $\mu_j \leftarrow h$ ).

Since the value of  $\phi$  is not known *a priori*, and even in scenarios where it is actually known beforehand, an additional

syndrome check is necessary for each guessed  $\phi$ , and the process usually must check all possible values of  $\phi$  anyway since more than one solution may exist.

#### D. The Completed Decoder

We are finally ready to state the full decoding method explicitly in Algorithm 1. It can be seen as a list decoding algorithm with possible failures. The polynomial decomposition of Equation 7 immediately suggests a simple and efficient way to compute the  $p$ -th roots needed at Step 9, namely, precompute  $r(x) \leftarrow \sqrt[p]{x} \bmod g(x)$  and  $r(x)^k \bmod g(x)$ , and then compute the  $p$ -th root of  $z(x) := \sum_k x^k z_k(x)^p$  as  $\sqrt[p]{z(x)} \bmod g(x) \leftarrow \sum_k r(x)^k z_k(x)$ . The test in Step 2 is unnecessary if  $g$  is irreducible. To find the zeroes of  $\sigma_\phi$  in Step 21 one can use the Chien search technique [23], in which case the multiplicities of each root can be determined as part of the search, or the Berlekamp trace algorithm [24].

#### V. CONCLUSION

We described a new decoding algorithm for square-free (in particular, irreducible) Goppa codes of degree  $t$  over  $\mathbb{F}_p$  that can correct  $(2/p)t$  errors in general, and up to  $t$  errors for certain distributions of error magnitudes of cryptographic interest. By attaining an correction capability of  $(2/3)t$  errors in characteristic 3 with high probability, our method outperforms the best previously known decoder for that case, and suggests that the corresponding average distance to the closest codeword is at least  $(4/3)t + 1$  for most irreducible ternary Goppa codes. Regardless of the characteristic, our proposal can correct a still larger number of errors that approaches (and probabilistically reaches)  $t$  as the distribution of error magnitudes becomes ever more skewed toward the predominance of some individual value. The method can be viewed as generalizing Patterson's binary decoding procedure, and is similarly efficient in practice.

A further increase in the number of correctable errors may be possible by resorting to list decoding and by extracting more information from the decoding process along the lines proposed by Bernstein [11]. This in principle might enable the correction of approximately  $n - \sqrt{n(n - (4/p)t)}$  errors in general, and possibly as many as  $n - \sqrt{n(n - 2t - 2)}$  errors depending on the distribution of error magnitudes. Furthermore, the ability to correct close to  $t$  errors with high probability means that smaller keys might be adopted for coding-based cryptosystems. Properly chosen parameters would keep the probability of decoding failure below the probability of breaking the resulting schemes by random guessing, while maintaining the security at the desired level. We leave the investigation of such possibilities for future research.

#### REFERENCES

- [1] D. J. Bernstein, J. Buchmann, and E. Dahmen, *Post-Quantum Cryptography*. Springer, 2008.
- [2] V. D. Goppa, "A new class of linear error correcting codes," *Problemy Peredachi Informatsii*, vol. 6, pp. 24–30, 1970.
- [3] K. K. Tzeng and C. R. P. Hartmann, "On the minimum distance of certain reversible cyclic codes," *IEEE Transactions on Information Theory*, vol. 16, no. 5, pp. 644–646, 1970.

#### Algorithm 1 Decoding $p$ -ary square-free Goppa codes

---

INPUT:  $\Gamma(L, g)$ , a Goppa code over  $\mathbb{F}_p$  where  $g$  is square-free.  
 INPUT:  $H \in \mathbb{F}_q^{r \times n}$ , a parity-check matrix in the form of Equation 1.  
 INPUT:  $c' = c + e \in \mathbb{F}_p^n$ , the received codeword with errors.  
 OUTPUT: a list of corrected codewords  $c \in \Gamma(L, g)$  ( $\emptyset$  upon failure).

```

1:  $s^\top \leftarrow Hc'^\top \in \mathbb{F}_q^r$ ,  $s_e(x) \leftarrow \sum_i s_i x^i$ .  $\triangleright$  N.B.  $Hc'^\top = He^\top$ .
2: if  $\nexists s_e^{-1}(x) \bmod g(x)$  then
3:   return  $\emptyset$ 
4: end if
5:  $S \leftarrow \emptyset$ 
6: for  $\phi \leftarrow 1$  to  $p - 1$  do  $\triangleright$  guess the correct scale factor  $\phi$ 
7:   for  $k \leftarrow 1$  to  $p - 1$  do
8:      $u_k(x) \leftarrow x^k + \phi k x^{k-1} / s_e(x) \bmod g(x)$ 
9:      $v_k(x) \leftarrow \sqrt[p]{u_k(x)} \bmod g(x)$ 
10:  end for
11:  Build the lattice basis  $A_\phi$  defined by Equation 9.
12:  Apply Algorithm 2 to reduce the basis of  $\Lambda(A_\phi)$ .
13:  for  $i \leftarrow 1$  to  $p$  do
14:    Let  $a$  denote the  $i$ -th row of the reduced basis of
     $\Lambda(A_\phi)$ 
15:    for  $j \leftarrow 0$  to  $p - 1$  do
16:      if  $\deg(a_j) > \lfloor (t - j)/p \rfloor$  then
17:        try next  $i \triangleright$  not a solution
18:      end if
19:    end for
20:     $\sigma_\phi(x) \leftarrow \sum_j x^j a_j(x)^p$ 
21:    Compute the set  $J$  such that  $\sigma_\phi(L_j) = 0$ ,  $\forall j \in J$ .
22:    for  $j \in J$  do
23:      Compute the multiplicity  $\mu_j$  of  $L_j$ .
24:       $e_j \leftarrow \phi \mu_j$ 
25:    end for
26:    if  $He^\top = s^\top$  then
27:       $S \leftarrow S \cup \{c' - e\}$ 
28:    end if
29:  end for
30:  return  $S$ 
31: end for
```

---

- [4] C. R. P. Hartmann and K. K. Tzeng, "Generalizations of the BCH bound," *Information and Control*, vol. 20, no. 5, pp. 489–498, 1972.
- [5] Y. Sugiyama, M. Kasahara, S. Hirasawa, and T. Namekawa, "Further results on Goppa codes and their applications to constructing efficient binary codes," *IEEE Transactions on Information Theory*, vol. 22, no. 5, pp. 518–526, 1976.
- [6] D. J. Bernstein, T. Lange, and C. Peters, "Wild McEliece," in *Selected areas in cryptography – SAC 2010*, ser. Lecture Notes in Computer Science, vol. 6544. Springer, 2011, pp. 143–158.
- [7] E. R. Berlekamp, *Algebraic Coding Theory*. New York, USA: McGraw-Hill, 1968.
- [8] Y. Sugiyama, M. Kasahara, S. Hirasawa, and T. Namekawa, "A method for solving key equation for decoding Goppa codes," *Information and Control*, vol. 27, no. 1, pp. 87–99, 1975.
- [9] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes*. North-Holland Mathematical Library, 1977, vol. 16.
- [10] V. Guruswami and M. Sudan, "Improved decoding of Reed-Solomon and algebraic-geometry codes," *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 1757–1767, 1999.
- [11] D. J. Bernstein, "List decoding for binary Goppa codes," in *Coding and cryptography – IWCC 2011*, ser. Lecture Notes in Computer Science, vol. 6639. Springer, 2011, pp. 62–80, <http://cr.yp.to/papers.html#goppalst>.

- [12] D. Augot, M. Barbier, and A. Couvreur, "List-decoding of binary Goppa codes up to the binary Johnson bound," in *Information Theory Workshop – ITW 2011*. IEEE, 2011, <http://arxiv.org/abs/1012.3439>.
- [13] R. McEliece, "A public-key cryptosystem based on algebraic coding theory," The Deep Space Network Progress Report, DSN PR 42-44, pp. 114–116, 1978, <http://ipnpr.jpl.nasa.gov/progressreport2/42-44/44N.PDF>.
- [14] E. Fujisaki and T. Okamoto, "Secure integration of asymmetric and symmetric encryption schemes," in *Advances in Cryptology – Crypto 1999*, ser. Lecture Notes in Computer Science, vol. 1666. Springer, 1999, pp. 537–554.
- [15] J.-C. Faugère, A. Otmani, L. Perret, and J.-P. Tillich, "Algebraic cryptanalysis of McEliece variants with compact keys," in *Advances in Cryptology – Eurocrypt 2010*, ser. Lecture Notes in Computer Science, vol. 6110. Springer, 2010, pp. 279–298.
- [16] C. Peters, "Information-set decoding for linear codes over  $\mathbb{F}_q$ ," in *Post-Quantum Cryptography Workshop – PQCrypto 2010*, ser. Lecture Notes in Computer Science, vol. 6061. Springer, 2010, pp. 81–94.
- [17] D. Micciancio, "The shortest vector problem is np-hard to approximate to within some constant," *SIAM Journal on Computing*, vol. 30, no. 6, pp. 2008–2035, 2001.
- [18] T. Mulders and A. Storjohann, "On lattice reduction for polynomial matrices," *Journal of Symbolic Computation*, vol. 35, pp. 377–401, 2003.
- [19] K. Lee and M. O'Sullivan, "List decoding of Reed-Solomon codes from a Gröbner basis perspective," *Journal of Symbolic Computation*, vol. 43, no. 9, pp. 645–658, 2008.
- [20] N. J. Patterson, "The algebraic decoding of Goppa codes," *IEEE Transactions on Information Theory*, vol. 21, no. 2, pp. 203–207, 1975.
- [21] R. Rivest, "All-or-nothing encryption and the package transform," in *Fast Software Encryption – FSE 1997*, ser. Lecture Notes in Computer Science, vol. 1267. Springer, 1997, pp. 210–218.
- [22] W. Bosma, J. Cannon, and C. Playoust, "The Magma algebra system. I. The user language," *J. Symbolic Comput.*, vol. 24, no. 3-4, pp. 235–265, 1997, computational algebra and number theory (London, 1993). [Online]. Available: <http://dx.doi.org/10.1006/jscs.1996.0125>
- [23] R. T. Chien, "Cyclic decoding procedure for the Bose-Chaudhuri-Hocquenghem codes," *IEEE Transactions on Information Theory*, vol. 10, no. 4, pp. 357–363, 1964.
- [24] E. R. Berlekamp, "Factoring polynomials over large finite fields," *Mathematics of Computation*, vol. 24, no. 111, pp. 713–715, 1970.

## APPENDIX A THE WEAK POPOV FORM

For ease of reference, we provide here a concise description of the Mulders-Storjohann polynomial lattice reduction algorithm based on the weak Popov form. We closely follow the exposition in [18], while attempting to make our description more implementation-friendly.

**Definition 1.** For  $1 \leq i \leq n$  the  $i$ -th pivot index vector  $I^M$  of a matrix  $M \in \mathbb{F}_q[x]^{n \times m}$  is defined as follows: if  $M_{ij} = 0$  for all  $1 \leq j \leq m$ , then  $I_i^M = 0$ , otherwise

- 1)  $\deg(M_{ij}) \leq \deg(M_{i,I_i^M})$  for  $1 \leq j < I_i^M$ ,
- 2)  $\deg(M_{ij}) < \deg(M_{i,I_i^M})$  for  $I_i^M < j \leq m$ .

**Definition 2.** The carrier set  $C^M$  of a matrix  $M \in \mathbb{F}_q[x]^{n \times m}$  is the set  $\{1 \leq i \leq n \mid I_i^M \neq 0\}$ . The  $i$ -th pivot element of  $M$ , denoted  $P_i^M$ , is the element  $P_i^M := M_{i,I_i^M}$  when  $I_i^M \neq 0$ , otherwise  $P_i^M := 0$ .

**Definition 3.** A matrix  $M \in \mathbb{F}_q[x]^{n \times m}$  is said to be in weak Popov form if the positive pivot indices of  $M$  are all different, i.e. if  $\forall k, \ell \in C^M : k \neq \ell \Rightarrow I_k^M \neq I_\ell^M$ .

The following theorem establishes that writing a matrix in weak Popov form yields short vectors in the lattice spanned by its rows.

**Theorem 3** ([18]). If matrix  $M \in \mathbb{F}_q[x]^{n \times m}$  is in weak Popov form and  $\ell$  is such that  $\deg(P_\ell^M) = \min_{1 \leq i \leq n} \{\deg(P_i^M)\}$ , then

all vectors in the  $\mathbb{F}_q[x]$ -module generated by the rows of  $M$  have degree at least  $\deg(P_\ell^M)$ .

*Proof:* See [18, Lemma 8.1]. ■

If  $k \in C^M$ ,  $\ell \neq k$  and  $\deg(M_{\ell,I_k^M}) \geq \deg(P_k^M)$ , there are unique  $c \in \mathbb{F}_q$  and  $e \in \mathbb{N}$  such that  $\deg(M_{\ell,I_k^M} - cx^e P_k^M) < \deg(M_{\ell,I_k^M})$ . In that case we call the operation  $M_\ell \leftarrow M_\ell - cx^e P_k^M$  the *simple transformation* of row  $k$  on row  $\ell$ . If  $I_\ell^M = I_k^M$ , the transformation is called of the *first kind*. Then an efficient algorithm to put a matrix in weak Popov form stems from the following observation:

**Theorem 4** ([18]).  $M \in \mathbb{F}_q[x]^{n \times m}$  is not in weak Popov form iff one can apply a simple transformation of the first kind on  $M$ , that is, not all non-zero pivot indices of  $M$  are different.

*Proof:* See [18, Lemma 2.1]. ■

Therefore, all one has to do to obtain the weak Popov form of a matrix  $M$  is to repeatedly check if  $M$  is already in the weak Popov form (by testing if all nonzero pivot indices are different) and, if it is not, apply a simple transformation of the first kind on it.

This process is summarised in Algorithm 2 for a matrix in the form of Equation 9, where  $n = m = p$ , the expected rank is  $r \leq p$ , and the degree of all rows is bounded by  $d \leq t$ . By Theorem 1, its complexity is  $O(p^3 t^2)$   $\mathbb{F}_q$  operations at most. Here  $\text{lead}(P)$  denotes the leading coefficient of  $P \in \mathbb{F}_q[x]$  and  $\text{rep}(I^A)$  denotes the number of occurrences of the most frequent value among the nonzero components of  $I^A$ , i.e.  $\text{rep}(I^A) := \max\{\# \{j \mid I_j^A = v\} \mid v \neq 0\}$ .

Written in this form, Algorithm 2 is strikingly similar to the modified Euclidean algorithm usually employed in the decoding of alternant codes [8], and actually coincides with that method for  $p = 2$ .

---

### Algorithm 2 Computing the weak Popov form

---

INPUT:  $A \in \mathbb{F}_q[x]^{p \times p}$  in the form of Equation 9.

OUTPUT: weak Popov form of  $A$ .

- 1:  $\triangleright$  Compute  $I^A$ :
  - 2: **for**  $j \leftarrow 1$  **to**  $p$  **do**
  - 3:    $I_j^A \leftarrow$  **if**  $\deg(A_{j,1}) > 0$  **then** 1 **else**  $j$
  - 4: **end for**
  - 5:  $\triangleright$  Put  $A$  in weak Popov form:
  - 6: **while**  $\text{rep}(I^A) > 1$  **do**
  - 7:   **for**  $k \leftarrow 1$  **to**  $p$  **such that**  $I_k^A \neq 0$  **do**
  - 8:     **for**  $\ell \leftarrow 1$  **to**  $p$  **such that**  $\ell \neq k$  **do**
  - 9:       **while**  $\deg(A_{\ell,I_k^A}) \geq \deg(A_{k,I_k^A})$  **do**
  - 10:           $c \leftarrow \text{lead}(A_{\ell,I_k^A}) / \text{lead}(A_{k,I_k^A})$
  - 11:           $e \leftarrow \deg(A_{\ell,I_k^A}) - \deg(A_{k,I_k^A})$
  - 12:           $A_\ell \leftarrow A_\ell - cx^e A_k$
  - 13:       **end while**
  - 14:      $\triangleright$  Update  $I_\ell^A$  and hence  $\text{rep}(I^A)$  if necessary:
  - 15:      $d \leftarrow \max\{\deg(A_{\ell,j}) \mid j = 1, \dots, p\}$
  - 16:      $I_\ell^A \leftarrow \max\{j \mid \deg(A_{\ell,j}) = d\}$
  - 17:   **end for**
  - 18: **end for**
  - 19: **end while**
  - 20: **return**  $A$
-

## APPENDIX B DECODING OTHER FAMILIES OF CODES?

For completeness, we briefly discuss whether and how one might attempt to use similar methods to decode a different family of alternant codes, including BCH codes and their permuted and/or punctured versions.

Let  $L \in \mathbb{F}_q^n$  be a sequence of  $n \leq q$  distinct nonzero elements, let  $D \in \mathbb{F}_q^n$  be a sequence of nonzero elements, and let  $H = \text{vdm}(L) \text{diag}(D)$ . For any word  $e \in \mathbb{F}_p^n$  we define the corresponding *alternant  $r$ -syndrome* polynomial  $s_e \in \mathbb{F}_q[x]$  to be  $s_e(x) := \sum_{i=0}^{r-1} s_i x^i$  where  $s^T := H e^T$ , i.e.

$$s_i = \sum_{j=0}^{n-1} e_j D_j L_j^i.$$

The alternant code  $\mathcal{A}(L, D, r)$  consists of the set  $\{e \in \mathbb{F}_p^n \mid s_e(x) \equiv 0\}$ .

Using the formula for the sum of a geometric sequence  $\sum_{i=0}^{r-1} u^i = (1-u^r)/(1-u)$  whereby  $\sum_{i=0}^{r-1} L_j^i x^i = (1-x^r L_j^r)/(1-xL_j) \equiv 1/(1-xL_j) \pmod{x^r}$ , one can see that

$$\begin{aligned} s_e(x) &= \sum_{i=0}^{r-1} \sum_{j=0}^{n-1} e_j D_j L_j^i x^i = \sum_{j=0}^{n-1} e_j D_j \sum_{i=0}^{r-1} L_j^i x^i \\ &\equiv \sum_{j=0}^{n-1} \frac{e_j D_j}{1-xL_j} \pmod{x^r}. \end{aligned}$$

The subfamily we will be interested in is that of alternant codes satisfying the restriction  $\xi_j := D_j/L_j \in \mathbb{F}_p \setminus \{0\}$  for all  $j$ , so that each value  $\xi_j$  can be lifted to  $\mathbb{Z}$  with a representative in range  $1 \dots p-1$ .

Let  $\phi \in \mathbb{F}_p \setminus \{0\}$  be a constant scalar. We define the generalized error locator polynomial for this family as

$$\sigma_\phi(x) := \prod_i (1 - xL_i)^{e_i \xi_i / \phi}. \quad (10)$$

The error positions are revealed by the *inverses* of the components of  $L$ , which are the roots of this polynomial. This definition coincides with the usual alternant error locator polynomial when  $p = 2$ , in which case  $D = L$  (hence, a permuted and/or punctured subcode of a binary BCH code).

Taking the derivative of the formal power series underlying  $\sigma_\phi$  in Equation 10, we get

$$\begin{aligned} \sigma'_\phi(x) &= \sum_j (e_j \xi_j / \phi) (1 - xL_j)^{e_j \xi_j / \phi - 1} (-L_j) \\ &\quad \prod_{i \neq j} (1 - xL_i)^{e_i \xi_i / \phi} \\ &= -(1/\phi) \sum_j \frac{e_j D_j}{1 - xL_j} \prod_i (1 - xL_i)^{e_i \xi_i / \phi} \\ &= -(1/\phi) \sigma_\phi(x) \sum_j \frac{e_j D_j}{1 - xL_j}, \end{aligned}$$

which over  $\mathbb{F}_q[x]$  reduces to

$$-\phi \sigma'_\phi(x) = \sigma_\phi(x) s_e(x) \pmod{x^r}. \quad (11)$$

This is the  $\phi$ -th key equation for this family of codes.

Now most of the techniques developed above for Goppa codes can be applied to solve Equation 11. The main difference

is that the error magnitudes are computed as a function of the multiplicity  $\mu_j$  of a root  $1/L_j$  of  $\sigma_\phi$  as  $e_j = \phi \mu_j / \xi_j$ .

Writing  $\sigma_\phi(x) = \sum_{k=0}^{p-1} x^k a_k(x)^p$  for some  $a_k(x)$  with  $\deg(a_k) \leq \lfloor (r-k)/p \rfloor$ , solutions to Equation 11 can be found as short vectors  $(a_0, a_1, \dots, a_{p-1})$  in the polynomial lattice spanned by the rows of the matrix

$$A_\phi = \begin{bmatrix} x^r & 0 & \dots & 0 \\ -v_1 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -v_{p-1} & 0 & \dots & 1 \end{bmatrix}$$

where the  $v_k(x)$  are polynomials satisfying  $v_k(x)^p = x^k - \phi k x^{k-1} / s_e(x) \pmod{x^r}$ , provided that these exist.

Here the major obstacle for this technique becomes apparent: inverting  $s_e(x) \pmod{x^r}$  is usually fine, but computing the  $v_k(x)$  polynomials is only very seldom possible. Specifically, assuming that  $x^k - \phi k x^{k-1} / s_e(x) \pmod{x^r}$  are uniformly distributed polynomials in  $\mathbb{F}_q[x]/x^r$  for a random code of this family, the probability that it is a  $p$ -th power mod  $x^r$  is only about  $(q^{r/p}/q^r)^{p-1} = p^{-mr(p-1)^2/p}$ , corresponding to the vanishing of all but a fraction  $1/p$  of the  $r$  coefficients of each of the  $p-1$  polynomials needed to build matrix  $A$ .

Therefore there is scant chance that this would work in practice, except possibly for some highly contrived code whose syndromes lead to suitable radicands with high probability. It is an open problem whether such codes exist and, if so, what they might look like.